

## AMENDMENTS TO THE CLAIMS

Please amend the claims as indicated hereafter.

### Claims:

1-75. (Cancelled)

76. (New) A method implemented by a television set-top terminal (STT) coupled via a bi-directional communication network to a server located remotely from said STT, said method comprising steps of:

receiving via a tuner in the STT a video presentation provided by the server;  
outputting by the STT at least a portion of the video presentation as a television signal;

receiving a first user input associated with a visual scene contained in the video presentation;

storing information related to said visual scene in a memory of the STT responsive to receiving the first user input;

outputting by the STT at least another portion of the video presentation as a television signal after the information has been stored in the memory of the STT;

receiving a second user input configured to request said visual scene in said video presentation after the STT has output the at least another portion of the video presentation; and

outputting by the STT a television signal comprising a portion of said video presentation starting from a location corresponding to said visual scene responsive to the second user input, wherein the location corresponding to said visual scene is identified by the STT using the information related to said visual scene.

77. (New) The method of claim 76, wherein storing information related to said visual scene in a memory of the STT includes storing information identifying a location of said visual scene in relation to a point in said video presentation other than a point corresponding to a beginning of an entirety of the video presentation.

78. (New) The method of claim 76, wherein the video presentation is a video-on-demand presentation, and wherein the server transmits the portion of said video presentation starting from said visual scene responsive to the second user input.

79. (New) The method of claim 76, further comprising the step of:

receiving a user input configured to assign a character sequence to said visual scene in said video presentation;

storing data corresponding to said character sequence in a memory of the STT responsive to receiving the user input configured to assign a character sequence; and

providing said character sequence simultaneously with an image corresponding to said visual scene responsive to subsequent user input.

80. (New) The method of claim 79, wherein said user input configured to assign a character sequence is received while said video presentation is being presented to said user.

81. (New) The method of claim 79, wherein said user input configured to assign a character sequence is received during a time period when the video presentation is not being provided by the STT.

82. (New) The method of claim 79, further comprising receiving a plurality of user inputs configured to assign a plurality of respective character sequences corresponding to a plurality of respective visual scenes that were bookmarked responsive to a plurality of respective user inputs, wherein the plurality of user inputs configured to assign the plurality of respective character sequences are received after the video presentation has been provided to the user.

83. (New) The method of claim 76, further comprising the step of:

receiving a user input configured to request information related to said visual scene in said video presentation; and  
providing the requested information responsive to receiving the user input configured to request information.

84. (New) The method of claim 76, wherein the first user input associated with the visual scene is received while the video presentation is being output by the STT in a normal playback mode, wherein outputting the video presentation by the STT is not interrupted responsive to the first user input.

85. (New) The method of claim 84, further comprising outputting information confirming that the visual scene has been bookmarked, wherein the information overlays a minority portion of a television screen being used to display the video presentation.

86. (New) The method of claim 85, wherein said information confirming that the visual scene has been bookmarked includes at least one of a banner and an icon.

87. (New) The method of claim 76, further comprising storing information related to said visual scene in a memory of the server responsive to receiving the first user input.

88. (New) The method of claim 76, wherein said first user input is received after display of the video presentation by a television coupled to the STT has been interrupted responsive to user input.

89. (New) The method of claim 76, wherein said second user input corresponds to a thumbnail image corresponding to the visual scene.

90. (New) The method of claim 76, wherein said visual scene is associated with a bookmark list associated with a plurality of visual scenes associated with a plurality of respective user inputs.

91. (New) The method of claim 76, further comprising associating a plurality of visual scenes with a plurality of respective bookmark lists associated with a plurality of respective users responsive to a plurality of respective user inputs.
92. (New) The method of claim 76, further comprising associating a plurality of visual scenes with a plurality of respective bookmark lists associated with a plurality of respective video presentations responsive to a plurality of respective user inputs.
93. (New) The method of claim 76, further comprising:  
after expiration of a rental access period corresponding to the video presentation,  
prompting said user to provide input indicating whether said information  
is to be deleted from the memory of the STT.
94. (New) The method of claim 76, further comprising:  
storing an image corresponding to said visual scene in a memory of the STT  
responsive to receiving the first user input;
95. (New) The method of claim 76, wherein said second user input requesting said visual scene corresponds to a thumbnail image corresponding to the visual scene, said thumbnail image being simultaneously provided with a plurality of thumbnail images corresponding to a plurality of visual scenes in the video presentation.

96. (New) A television set-top terminal (STT) coupled via a bi-directional communication network to a server located remotely from said STT, said STT comprising:
- a tuner configured to receive a video presentation provided by the server;
  - memory;
  - a processor that is programmed to enable the STT to:
    - output at least a portion of the video presentation as a television signal;
    - store information related to a visual scene contained in the video presentation in the memory responsive to the STT receiving a first user input associated with said visual scene;
    - output at least another portion of the video presentation as a television signal after the information has been stored in the memory;
    - output responsive to the STT receiving a second user input a television signal comprising a portion of said video presentation starting from a location corresponding to said visual scene;
  - wherein the location corresponding to said visual scene is identified by the STT using the information related to said visual scene; and
  - wherein the television signal comprising the portion of said video presentation starting from a location corresponding to said visual scene is output after the at least another portion of the video presentation is output as a television signal.
97. (New) The STT of claim 96, wherein said visual scene is associated with a bookmark list associated with a plurality of visual scenes corresponding to a plurality of respective user inputs.
98. (New) The STT of claim 96, wherein the processor is programmed to associate a plurality of visual scenes with a plurality of respective bookmark lists associated with a plurality of respective users responsive to a plurality of respective user inputs.
99. (New) The STT of claim 96, wherein the processor is programmed to associate a plurality of visual scenes with a plurality of respective bookmark lists associated with a plurality of respective video presentations responsive to a plurality of respective user inputs.

100. (New) The STT of claim 96, wherein the processor is configured to prompt said user to provide input indicating whether said data is to be deleted from the memory of the STT.

101. (New) The STT of claim 96, wherein the processor is configured to enable the STT to store in the memory an image corresponding to said visual scene responsive to receiving the first user input.

102. (New) A method implemented by a television set-top terminal (STT) coupled via a bi-directional communication network to a server located remotely from said STT, said method comprising steps of:

providing a plurality of images corresponding to a plurality of locations in a video presentation, the video presentation being received by the STT from the server via the bi-directional communication network, wherein each of the plurality of locations is associated with a respective user input received by the STT; and

providing a plurality of names corresponding to the plurality of images, wherein each of the plurality of names was selected by a respective user input received by the STT.

103. (New) The method of claim 102, wherein at least one of the plurality of locations was identified by a respective user input while the video presentation was being output by the STT in a normal play mode.

104. (New) The method of claim 102, wherein at least one of the plurality of locations was identified by a respective user input while the video presentation was not being output by the STT.

105. (New) The method of claim 102, wherein at least one of the plurality of names was selected by a respective user input from a list of names corresponding to one of the plurality of images.

106. (New) A television set-top terminal (STT) coupled via a bi-directional communication network to a server located remotely from said STT, said STT comprising:

a processor programmed to enable the STT to output a plurality of images and a plurality of corresponding names, the plurality of images corresponding to a plurality of locations in a video presentation, the video presentation being received by the STT from the server via the bi-directional communication network, wherein each of the plurality of locations was identified by a respective user input received by the STT, and wherein each of the plurality of names was selected by a respective user input received by the STT.

107. (New) The STT of claim 106, wherein at least one of the plurality of locations was identified by a respective user input while the video presentation was being output by the STT in a normal play mode.

108. (New) The STT of claim 106, wherein at least one of the plurality of locations was identified by a respective user input while the video presentation was not being output by the STT.

109. (New) The STT of claim 106, wherein at least one of the plurality of names was selected by a respective user input from a list of names corresponding to one of the plurality of images.

110. (New) A method implemented by a television set-top terminal (STT) coupled via a bi-directional communication network to a server located remotely from said STT, said method comprising steps of:

identifying by the STT a plurality of locations in a video presentation responsive to a plurality of respective user inputs, the video presentation being received by the STT from the server via the bi-directional communication network;

associating by the STT a plurality of respective names with the plurality of locations responsive to a plurality of respective user inputs, wherein the plurality of respective names include a first name and a second name, and wherein the plurality of locations include a first location and a second location;

outputting by the STT a first television signal configured to encode the first name and an image corresponding to the first location;

outputting by the STT a second television signal responsive to user input received while the first television signal was being output by the STT, the second television signal being configured to encode the second name and an image corresponding to the second location.

111. (New) The method of claim 110, further comprising:

receiving a user input corresponding to the second image; and

providing a portion of the video presentation starting from a location corresponding to the second image, responsive to receiving the user input corresponding to the second image.



112. (New) A method implemented by a television set-top terminal (STT) coupled via a bi-directional communication network to a server located remotely from said STT, said method comprising steps of:

- identifying a plurality of locations in a video presentation responsive to a plurality of respective user inputs, the video presentation being received by the STT from the server via the bi-directional communication network;
- associating a plurality of respective names with the plurality of locations responsive to a plurality of respective user inputs;
- providing a list that includes the plurality of names;
- receiving user input corresponding to one of the plurality of names included in the list; and
- providing a portion of the video presentation starting from a location corresponding to said one of the plurality of names.

113. (New) The method of claim 112, wherein at least one of the plurality of locations was identified by a respective user input while the video presentation was being output by the STT in a normal play mode.

114. (New) The method of claim 112, wherein at least one of the plurality of names was selected by a respective user input from a list of names provided by the STT.

115. (New) A method implemented by a television set-top terminal (STT) coupled via a bi-directional communication network to a server located remotely from said STT, said method comprising:

- receiving via a tuner in the STT a video presentation provided by the server;
- outputting by the STT at least a portion of the video presentation as a television signal;
- receiving a first user input associated with a visual scene contained in the video presentation;
- storing information related to said visual scene in a memory of the STT responsive to receiving the first user input;

outputting by the STT at least another portion of the video presentation as a television signal after the information has been stored in the memory of the STT;

receiving a second user input configured to request said visual scene in said video presentation after the STT has output the at least another portion of the video presentation; and

outputting by the STT a television signal comprising a portion of said video presentation starting from a location corresponding to said visual scene responsive to the second user input, wherein the location corresponding to said visual scene is identified by the STT using the information related to said visual scene;

receiving user input configured to assign a character sequence to said visual scene in said video presentation;

storing data corresponding to said character sequence in a memory of the STT responsive to receiving the user input configured to assign a character sequence;

providing said character sequence simultaneously with an image corresponding to said visual scene responsive to user input;

receiving a user input configured to request information related to said visual scene in said video presentation;

providing the requested information responsive to receiving the user input configured to request information;

outputting information confirming that the visual scene has been bookmarked; wherein the information overlays a minority portion of a television screen being used to display the video presentation;

wherein said information confirming that the visual scene has been bookmarked includes at least one of a banner and an icon;

wherein the video presentation is a video-on-demand presentation;

wherein the server transmits the portion of said video presentation starting from said visual scene responsive to the second user input;

wherein the first user input associated with the visual scene is received while the

video presentation is being output by the STT in a normal playback mode;  
and  
wherein outputting the video presentation by the STT is not interrupted  
responsive to the first user input.